PATENT
Attorney Docket 2676-6388US

NOTICE OF EXPRESS MAILING

Express Mail Mailing Label Number: EV326919920US

Date of Deposit with USPS: March 29, 2004

Person making Deposit: Christopher Haughton

APPLICATION FOR LETTERS PATENT

for

AN EFFICIENT SYSTEM FOR RNA SILENCING

Inventors:
Anna Depicker
Helena van Houdt

Attorney:
Allen C. Turner
Registration No. 33,041
TRASKBRITT
P.O. Box 2550
Salt Lake City, Utah 84110
(801) 532-1922

TITLE OF THE INVENTION

AN EFFICIENT SYSTEM FOR RNA SILENCING

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application is a continuation of PCT International Patent Application No. PCT/EP/02/11188, filed on October 2, 2002, designating the United States of America, and published, in English, as PCT International Publication No. WO 03/031632 A1 on April 17, 2003, the contents of the entirety of which is incorporated by this reference.

TECHNICAL FIELD

[0002] The invention relates generally to biotechnology, and more particularly to a method for efficient RNA silencing in eucaryotic cells, particularly plant cells. Consequently, the method can be used to reduce the phenotypic expression of an endogenous gene in a plant cell. Furthermore, the method can be applied in a high throughput screening for RNA silencing.

BACKGROUND

[0003] "RNA silencing" is a type of gene regulation based on sequence-specific targeting and degradation of RNA. The term encompasses related pathways found in a broad range of eukaryotic organisms, including fungi, plants, and animals.

[0004] In plants, RNA silencing serves as an antiviral defense and many plant viruses encode suppressors of silencing. Also, it becomes clear that elements of the RNA silencing system are essential for gene regulation in development. The emerging view is that RNA silencing is part of a sophisticated network of interconnected pathways for cellular defense, transposon surveillance, and regulation of development. Based on the sequence specific RNA degradation, RNA silencing has become a powerful tool to manipulate gene expression experimentally. RNA silencing was first discovered in transgenic plants, where it was termed co-suppression or posttranscriptional gene silencing (PTGS). Sequence-specific RNA degradation processes related to PTGS have also been found in ciliates, fungi, and a variety of animals from *Caenorhabditis elegans* to mice (RNA interference).

[0005] A key feature uniting the RNA silencing pathways in different organisms is the importance of double-stranded RNA (dsRNA) as a trigger or an intermediate. The dsRNA is

cleaved into small interfering RNAs (21 to 25 nucleotides) of both polarities, and these are thought to act as guides to direct the RNA degradation machinery to the target RNAs. An intriguing aspect of RNA silencing in plants is that it can be triggered locally and then spread via a mobile silencing signal. In plants, RNA silencing is correlated with methylation of homologous transgene DNA in the nucleus. Other types of epigenetic modifications may be associated with silencing in other organisms.

[0006] It is known from the art that transgenes encoding ds or self-complementary (hairpin) RNAs of endogenous gene sequences are highly effective at directing the cell's degradation mechanism against endogenous (ss) mRNAs, thus giving targeted gene suppression. This discovery has enabled the transgenic enhancement of a plant's defense mechanism against viruses that it is unable to combat unaided. It has also shed light on how antisense and co-suppression might operate: by the inadvertent integration of two copies of the transgenes in an inverted repeat orientation, such that read-through transcription from one gene into the adjacent copy produces RNA with self-complementary sequences.

[0007] RNA silencing is induced in plants by transgenes designed to produce either sense or antisense transcripts. Furthermore, transgenes engineered to produce self-complementary transcripts (dsRNAs) are potent and consistent inducers of RNA silencing. Finally, replication of plant viruses, many of which produce dsRNA replication intermediates, causes a type of RNA silencing called Virus Induced Gene Silencing (VIGS). Whether VIGS, and the different types of transgene-induced RNA silencing in plants result from similar or distinct mechanisms is still a matter of debate. However, recent genetic evidence raises the possibility that the RNA silencing pathway is branched and that the branches converge in the production of dsRNA.

SUMMARY OF THE INVENTION

[0008] Until recently, RNA silencing was viewed primarily as a thorn in the side of plant molecular geneticists, limiting expression of transgenes and interfering with a number of applications that require consistent, high-level transgene expression. With our present understanding of the process, however, it is clear that RNA silencing could have enormous potential for engineering control of gene expression, as well as for the use as a tool in functional genomics. It could be experimentally induced and targeted to a single specific gene or even to a

family of related genes. Likewise, ds RNA-induced TGS may have similar potential to control gene expression. Although methods for RNA silencing have been described in the art (*e.g.*, WO99/53050, WO99/32619, WO99/61632, and W098/53083), a need exists to develop alternative and more efficient tools for RNA silencing.

[0009] In the present invention, we have developed a highly efficient method for RNA silencing that can also be used as a tool for high throughput silencing. The method uses a host that carries already a silenced locus and a second recombinant gene comprising a region that is homologous with the silenced locus. Although it is known that the recombinant gene will be silenced, we have surprisingly found that also target genes, which have no significant homology with the silenced locus but have homology with the recombinant gene, are efficiently silenced.

[0010] The present invention deals with an efficient method for RNA silencing in a eucaryotic host. The method makes use of a host that already comprises a silenced locus. Such a silenced locus can for example be generated by methods known in the art. For example the publication of De Buck and Depicker, 2001 and other publications, and also PCT patent publications WO99/53050, WO99/32619, WO99/61632, and W098/53083 describe methods to obtain RNA silencing and for generating a silenced recombinant locus. The 'target gene' is here defined as the gene of interest for silencing or to down-regulate its expression. An important aspect of this invention is that the target gene has no significant homology with the silenced locus. No significant homology means that either the overall homology is less than 40, 35, 30, 25% or even less or that no contiguous stretch of at least 23 identical nucleotides are present (Thomas et al., 2001). Homology is typically measured using sequence analysis software (*e.g.*, Sequence Analysis Software Package of the Genetics Computer Group, University of Wisconsin Biotechnology Center, 1710 University Avenue, Madison, Wis. 53705). Such software matches similar sequences by assigning degrees of homology to various insertions, deletions, substitutions, and other modifications. Silencing of the target gene in the present invention occurs via an intermediate step and hence our method is designated as domino silencing (FIG. 1). In the intermediate step a recombinant gene construct is introduced by transformation into the host comprising the silenced locus. The recombinant gene construct has a region of homology with the silenced locus already present. The region of homology is preferably more than 60, 70, 80, 90, 95 or even more than 99% homologous. The homologous region between the silenced locus and the recombinant gene can be found in the 5' untranslated or 3' untranslated region of

the recombinant gene construct. Furthermore, the recombinant gene construct has a region of minimal 23 nucleotides (Thomas et al., 2001), but preferably longer, that are identical with the target gene, or has a region of overall homology of more than 60, 70, 80, 90, 95 or even more than 99%. A recombinant gene is defined herein as a construct which does not naturally occur in nature. A non-limiting example of a recombinant gene construct is a construct wherein the coding region of a gene is operably linked to a 5' untranslated region and/or to a 3' untranslated region of one or more other genes, alternatively the 5' or 3' untranslated region is an artificial sequence.

[0011] Thus, in one embodiment the invention provides a method for obtaining efficient RNA silencing of a target gene comprising the introduction of a recombinant gene into a host that comprises a silenced locus and an unsilenced target gene whereby the recombinant gene comprises a region that is homologous with the silenced locus and whereby the target gene has homology with the recombinant gene but has no significant homology with the silenced locus.

[0012] In another embodiment, the method is used wherein the host is a plant or plant cell.

[0013] In another embodiment, the method of the invention can be used for high throughput gene silencing. Indeed, a recombinant gene library can be made wherein for example every gene or coding region thereof is combined with (operably linked with) a region of homology with the silenced gene that resides in the silenced locus and the recombinant gene library can be transformed to an eukaryotic host or individual (specific) genes derived from the recombinant gene library can be transformed into an eukaryotic host wherein silencing of specific genes is wanted.

[0014] In yet another embodiment, the invention provides a plant or plant cell that comprises a silenced locus and wherein a silenced target gene is obtained through the introduction of a recombinant gene according to the current method of the invention.

[0015] In yet another embodiment, the RNA silencing of the target gene is obtained in more than 80, 85, 90 or 95% of the transgenic organisms.

[0016] In yet another embodiment, the RNA silencing of the target gene occurs at an efficiency of more than 80, 85, 90 or 95 % as compared to the level of the unsilenced expression of the target gene.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

[0017] FIG. 1: Schematic outline of homology between a silenced locus X, a recombinant gene Y and a target gene Z.

[0018] FIG. 2: Schematic outline of the T-DNA constructs that are present in silenced locus X_1, recombinant gene Y_1 and target gene Z_1 (T-DNAs of pGVCHS287, pGUSchsS and pXD610 respectively) and of the transcript homology between X_1, Y_1 and Z_1.

[0019] LB and RB: left and right T-DNA border respectively; Pnos: nopaline synthase promoter; hpt: hygromycin phosphotransferase coding sequence; 3'nos: 3'untranslated region of the nopaline synthase gene; P35S; Cauliflower mosaic virus 35S promoter; nptII c.s., neomycin phosphotransferase II coding sequence; 3'chs: 3'untranslated region of the chalcone synthase gene of Anthirrinum majus; +1: transcription start; A_n: poly A-tail; gus c.s.: β-glucuronidase coding sequence; Pss: promoter of the small subunit of rubisco; bar: phosphinotricine transferase coding sequence; 3'g7: 3'untranslated region of the Agrobacterium octopine T-DNA gene 7; 3'ocs: 3'untranslated region of octopine synthase gene.

[0020] FIG. 3: Schematic outline of the T-DNA construct present in silenced locus X_1 and of the transiently introduced T-DNAs Y_2 (T-DNAs of pGVCHS287 and pPs35SCAT1S3chs, respectively) and of the transcript homology between X_1, Y_2 and Z_2 (the catalase1 endogene). Abbreviations as in FIG. 2

[0021] FIG. 4: Schematic outline of the T-DNA constructs present in silenced locus X_2 and of the transiently introduced T-DNAs Y_2 (T-DNAs of pGUSchsS + pGUSchsAS, and pPs35SCAT1S3chs, respectively) and of the transcript homology between X_2, Y_2 and Z_2 (the catalase1 endogene). Abbreviations as in FIG. 2

[0022] FIG.5: pPs35SCAT1S3chs

DETAILED DESCRIPTION OF THE INVENTION

[0023] A post-transcriptionally silenced inverted repeat transgene locus can trigger silencing of a reporter gene producing non-homologous transcripts.

[0024] We studied the interaction between three transgene loci X_1, Y_1 and Z_1 (FIG. 2. For a detailed description of all loci and constructs, see materials and methods) to address the question whether or not a stepwise homology between loci can lead to silencing.

[0025] It has been demonstrated previously that the post-transcriptionally silenced *nptII* genes in locus X_1 are capable to in trans silence transiently expressed genes with partial transcript homology to their *nptII* transcripts (Van Houdt et al., 2000 b). We subsequently found that also a stably expressed β-glucuronidase (*gus*) gene (in locus Y_1), with partial transcript homology to the *nptII* transcripts of the silencing inducing locus X_1, becomes efficiently silenced in trans (FIG. 2: X_1 and Y_1 and table 1: X_1Y_1 compared to Y_1). On the contrary, the *nptII* genes of locus X_1 are not able to trigger silencing of the gus genes in locus Z_1 which is expected as the genes of both loci produce transcripts without significant homology (FIG. 2). The homology between the two transcripts of X_1 and Y_1 is mainly situated in the 3'untranslated region (250 nucleotides), but also the 5'untranslated sequences show a small region of homology (29 nucleotides). These results demonstrate that the in trans silencing effects are not triggered by promoter homology. When Y_1 and Z_1 loci are combined in so called Y_1Z_1 hybrids both types of gus genes, having transcript homology in the gus coding sequence of 1809 nucleotides, remain highly expressed as reflected in the normal gus activity showing that the RNA silencing mechanism does not become activated (Table 1: Y_1Z_1 compared to Y_1 and Z_1). Surprisingly, upon creation of a stepwise homology between X_1 and Z_1 by introducing locus Y_1, the new observation described here is that also the gus expression in locus Z_1 is reduced in $X_1Y_1Z_1$ plants (Table 1: $X_1Y_1Z_1$ compared to Y_1Z_1). Thus, creating a stepwise homology between a silenced locus and a target gene by introducing a recombinant gene is sufficient to trigger silencing of the target.

[0026] Silencing inducing transgene loci can trigger silencing of a non-homologous endogene.

[0027] We further assessed the universality and the usefulness in high throughput functional gene analyses of silencing elicited by a stepwise homology in trans, called domino silencing. Therefore, we evaluated whether the expression of the tobacco endogenous catalase1 (cat1) genes is reduced in plants carrying a silencing locus (X locus) showing no significant homology with the catalase endogene by introducing a recombinant gene (Y construct). As silencing locus we used either X_1 or X_2 (FIG. 2: locus X_1, FIG. 3: locus X_2), in either case containing the 3' chalcone synthase sequences of Anthirrinum majus (3'chs). As transmitter for silencing we constructed a recombinant gene composed of the catalase1 coding sequence and the 3' chs region under control of the 35S promoter (P35S) (residing on T-DNA pPs35SCAT1S3chs,

Figs. 2 and 3: T-DNA in Y_2). The recombinant cat1 3'chs genes (Y2) were introduced in tobacco leaves bearing locus X_1 (or X_2) via Agrobacterium injection. As a negative control, we introduced a recombinant gene in which the cat1 coding sequence is replaced by the gus coding sequence (pGUSchsS, T-DNA construct as in locus Y_1 FIG. 1). In this case, no stepwise homology is created between the silencing inducing locus and the target catalase endogenes. As a positive control, the recombinant construct Y_2 was also introduced in transgenic tobacco with silenced catalase1 genes by the presence of a catalase1 antisense construct (Cat1AS in Champnongpol et al., 1996). Sixteen days after Agrobacterium injection, the catalase activity was determined in protein extracts of injected leaf tissue and compared with the activity in non-injected wild type (SR1) leaf tissue (Table 2). The results indicate that domino silencing is also applicable to endogenes since the catalase activity is clearly reduced in 6 out of 7 samples, while it remains high in the negative controls. In conclusion, not only an inverted repeat-bearing silencing-inducing transgene locus, but also a silencing-inducing locus in which the two residing chimeric genes give rise to transcripts with complementarity in the 3'UTR (3'chs)(FIG. 3: X_2), is able to trigger domino silencing reducing endogenous catalase expression.

[0028] Table 1: Results of a GUS-activity determination in protein extracts of leaf tissue harvested from tobacco plants containing different combinations of the loci X_1, Y_1 and Z_1 (FIG. 2). The mean values of a number of plants (n) are given.

Genotype	GUS-act.[1] 4 weeks[2] U GUS/mg TSP	N	GUS-act. Mature[2] U GUS/mg TSP	n
X_1	<[3]	1	<	1
Y_1	368 ± 165[4]	9	n.d.	-
Z_1	126 ± 30	10	48 ± 8	5
X_1Y_1	2 ± 1	4	4 ± 2	4
X_1Z_1	139 ± 35	9	46 ± 14	5
Y_1Z_1	477 ± 101	10	231 ± 106	6
$X_1Y_1Z_1$[5] → Y_1Z_1	195 ± 104	16	315 ± 46	8
→ $X_1Y_1Z_1$	4 ± 3	22	12 ± 4	9

[1] The mean GUS-activity (GUS-act.) was calculated, using n samples and expressed as units (U) GUS per milligram of total soluble protein (TSP).

[2] The plants were analyzed in two different developmental stages; 4 weeks after sowing and at a mature stage just before onset of flowering.

[3] below detection limit (1 U GUS/mg TSP)

[4] standard deviation

[5] Growth of $X_1Y_1Z_1$ plants was performed in conditions that both Y_1Z_1 and $X_1Y_1Z_1$ plants were able to develop. A PCR screen with X_1-specific primers was performed to discriminate between presence and absence of X_1.

n.d. not determined

[0029] Table 2: Results of a catalase-activity determination in protein extracts of leaf tissue harvested from Agrobacterium injected tobacco leaves.

Genotype injected Plant	Construct introduced via Agrobacterium injection	catalase activity 16 days after injection (60 μg TSP)
WT (SR1)	- (non-injected)	-0.2116[2] 100%[3]
X_1	PGUSchsS	-0.2556 121%
X_1	Y_2	-0.0589 27%
X_1[4]	Y_2	-0.0698 33%
X_2	PGUSchsS	-0.1782 84%
X_2	Y_2	-0.0641 30%
X_2	Y_2	-0.0987 47%
X_2[4]	Y_2	-0.0914 43%
X_2[4]	Y_2	-0.1996 94%
X_2[4]	Y_2	-0.0627 30%
Cat1AS	Y_2	-0.0439 21%

[1] X_1, see FIG. 3; X_2, see FIG. 4.

[2] The mean of two samples independently measured (-0.2270 and –0.1963).

[3] The catalase activity in wild type SR1 tobacco leaves was set to 100%.

[4] 24 hours after Agrobacterium injection, the plants were placed under high light conditions for 24 hours (1000 μmol / m^2 s). This treatment is known to stimulate endogenous catalase 1 transcription. As the degree of cat suppression is similar in uninduced as in induced situation, the data indicate that enhanced transcription of the endogenous catalase target is not required to trigger domino silencing.

EXAMPLES

Materials and Methods

Plasmid construction

[0030] pPs35SCAT1S3chs: The T-DNA of this plasmid is schematically shown in FIG. 3:Y_2 and the nucleotide sequence is depicted in SEQ ID NO:1 of the accompanying and incorporated herein SEQUENCE LISTING.

Description of the transgene loci and production of hybrid plants

[0031] Locus X_1 harbors an inverted repeat about the right T-DNA border of construct pGVCHS287, carrying a neomycin phosphotransferase II (*nptII*) gene under the control of the Cauliflower mosaic virus 35S promoter (P35S) and the 3'signalling sequences of the Anthirrinum majus chalcone synthase gene (3'chs). The *nptII* genes are post-transcriptionally silenced and can trigger in trans silencing and methylation of homologous target genes (Van Houdt et al., 2000 a and b and FIG. 2).

[0032] Locus Y_1 contains a single copy of the pGUSchsS T-DNA, containing a gus gene under the control of P35S and 3'chs (in transformant GUSchsS29) and shows normal levels of gus expression (FIG. 2).

[0033] Locus Z_1 contains more than one copy of the pXD610 T-DNA, harboring the gus gene under control of P35S and the 3'untranslated region (UTR) of the nopaline synthase gene (3'nos), (in plant LXD610-2) and shows normal gus expression (De Loose et al., 1995 and FIG. 2).

[0034] Locus X_2 contains a single copy of both the pGUSchsS and pGUSchsAS T-DNA (in transformant GUSchsS+GUSchsAS 11) and triggers silencing in cis of the gus genes, but also in trans of (partially) homologous genes (FIG. 4).

[0035] X_1 and Z_1 hemizygous plants were obtained as hybrid progeny of the crossing of tobacco plants homozygous for locus X_1 (=Holo1; Van Houdt et al., 2000 a and b) and homozygous for locus Z_1 (=LXD610-2/9 De Loose et al., 1995) to wild type SR1 respectively. Y_1 hemizygous plants were obtained by crossing the hemizygous primary tobacco transformant GUSchsS29 to SR1 and selecting for the presence of locus Y_1 in the hybrid progeny. X_1Y_1 and Y_1Z_1 hemizygous plants are the hybrid progeny plants of the cross between Holo1 and GUSchsS29 and between GUSchsS29 and LXD610-2/9 respectively that are selected for the

presence of Y_1. X_1Z_1 hemizygous plants are the hybrid progeny of the cross between Holo1 and LXD610-2/9. $X_1Y_1Z_1$ hemizygous plants were obtained by crossing X_1Y_1 hemizygous plants to LXD610-2/9; as we only selected for the presence of Y_1 in the hybrid progeny both Y_1Z_1 and $X_1Y_1Z_1$ hemizygous plants were obtained.

Preparation of Agrobacteria and injection

[0036] The Agrobacteria C58C1Rif^R (pGV2260) (pGUSchsS)Cb^R,PPT^R or C58C1Rif^R(pMP90) (pPs35SCAT1S3chs)Gm^R,PPT^R were mainly grown as described by Kapila et al., 1997 except that the Agrobacteria were resuspended in MMA to a final OD_{600} of 1. Greenhouse grown plants of 10 to 15 cm in height were used. Half of the third top leaf was injected via the lower surface using a 5ml syringe while the leaf remained attached to the plant. The plants were kept in the greenhouse and 16 days after injection three to four discs of 11 mm in diameter were excised from the injected tissue for the preparation of a fresh protein extract to determine the catalase activity.

Enzymatic assays

[0037] Preparation of the protein extracts and GUS-activity measurements were done as previously described (Van Houdt et al., 2000 b). Preparation of the protein extracts for catalase-activity measurement and the spectrophotometric catalase-activity determination was done according to Champnongpol et al., 1996.

References

Van Houdt, H., Kovarik, A., Van Montagu, M., and Depicker, A. (2000 a). Cross-talk between posttranscriptionally silenced neomycin phosphotransferase II transgenes. *FEBS Lett.* 467, 41-46.

Van Houdt, H., Kovarik, A., Van Montagu, M., and Depicker, A. (2000 b) Both sense and antisense RNAs are targets for the sense transgene-induced posttranscriptional silencing mechanism. *Mol. Gen. Genet.* 263, 995-1002.

De Loose, M., Danthinne, X., Van Bockstaele, E., Van Montagu, M. and Depicker, A., (1995) Different 5'leader sequences modulate β-glucuronidase accumulation levels in transgenic *Nicotiana tobacum* plants. *Euphytica* 85, 209-216.

Kapila, J., De Rycke, R., Van Montagu, M. and Angenon, G. (1997) An *Agrobacterium*-mediated transient gene expression system for intact leaves. *Plant Science* 122, 101-108.

Champnongpol, S., Willekens, H., Langebartels, C., Van Montagu, M., Inzé, D., and Van Camp, W. (1996) Transgenic tobacco with a reduced catalase activity develops necrotic lesions and induces pathogenesis-related expression under high light. *Plant J.* 10(3), 491-503.

Thomas, C. L., Jones, L., Baulcombe, D.C. and Maule, A.J. (2001) Size constraints for targetting post-transcriptional gene silencing and for RNA-directed methylation in *Nicotiana benthamiana* using potato virus X vector. *Plant J.* 25(4), 417-425.

De Buck, S. and Depicker, A. (2001) Disruption of their palindromic arrangement leads to selective loss of DNA methylation in inversely repeated gus transgenes in Arabidopsis. *Mol. Gen. Genom.* 265, 1060-1068.

SEQUENCE LISTING

<110> VLAAMS INTERUNIVERSITAIR INSTITUUT VOOR BIOTECHNOLOGIE VZW

<120> An efficient system for RNA silencing

<130> ADP/Dom/V097

<150> EP01203760.2

<151> 2001-10-05

<160> 1

<170> PatentIn version 3.1

<210> 1

<211> 10635

<212> DNA

<213> Artificial Sequence

<220>

<223> pPs35SCAT1S3chs

<400> 1

agattcgaag ctcggtcccg tgggtgttct gtcgtctcgt tgtacaacga aatccattcc 60

cattccgcgc tcaagatggc ttcccctcgg cagttcatca gggctaaatc aatctagccg 120

acttgtccgg tgaaatgggc tgcactccaa cagaaacaat caaacaaaca tacacagcga 180

cttattcaca cgcgacaaat tacaacggta tatatcctgc cagtactcgg ccgtcgaata 240

acttcgtata atgtatgcta tacgaagtta tgaattcgcg ctctatcata gatgtcgcta 300

taaacctatt cagcacaata tattgttttc attttaatat tgtacatata agtagtaggg 360

tacaatcagt aaattgaacg gagaatatta ttcataaaaa tacgatagta acgggtgata 420

tattcattag aatgaaccga aaccggcggt aaggatctga gctacacatg ctcaggtttt 480

ttacaacgtg cacaacagaa ttgaaagcaa atatcatgcg atcataggcg tctcgcatat 540

ctcattaaag cagctggaag atttgatgga tcctcatcag atctcggtga cgggcaggac 600

cggacggggc ggtaccggca ggctgaagtc cagctgccag aaacccacgt catgccagtt 660

cccgtgcttg aagccggccg cccgcagcat gccgcggggg gcatatccga gcgcctcgtg 720

catgcgcacg ctcgggtcgt tgggcagccc gatgacagcg accacgctct tgaagccctg 780

tgcctccagg gacttcagca ggtgggtgta gagcgtggag cccagtcccg tccgctggtg 840

gcgggggggag acgtacacgg tcgactcggc cgtccagtcg taggcgttgc gtgccttcca 900

ggggcccgcg taggcgatgc cggcgacctc gccgtccacc tcggcgacga gccagggata 960

gcgctcccgc agacggacga ggtcgtccgt ccactcctgc ggttcctgcg gctcggtacg 1020

gaagttgacc gtgcttgtct cgatgtagtg gttgacgatg gtgcagaccg ccggcatgtc 1080

cgcctcggtg gcacggcgga tgtcggccgg gcgtcgttct gggctcatgg tagatctgtt 1140

taaacgttaa cggattgaga gtgaatatga gactctaatt ggataccgag gggaatttat 1200

ggaacgtcag tggagcattt ttgacaagaa atatttgcta gctgatagtg accttaggcg 1260

acttttgaac gcgcaataat ggtttctgac gtatgtgctt agctcattaa actccagaaa 1320

cccgcggctg agtggctcct tcaatcgttg cggttctgtc agttccaaac gtaaaacggc 1380

ttgtcccgcg tcatcggcgg gggtcataac gtgactccct taattctccg ctcatgatca 1440

agctacctca gcaggatccg gcgcgccatg gtcgataaga aaaggcaatt tgtagatgtt 1500

aattcataac atctcctcca tgacttaaaa aacttgcaaa agatttatat agaaatactt 1560

aaatattttg actaaaaaaa aaaaaaaaaa aacacacaca taaaccaaca aataacataa 1620

attatttta tatagccttt atttcaatga tcacaacgaa acaatacaag tacaaagcgt 1680

tacaagagag aaatcgccaa tatagctcac atgcagcaca catcacaata ataggtaacc 1740

atgtccactt ttttattacg gaaataagaa aataacccaa ccccgtacc cgggttcata 1800

tgcttggtct cacattaagc ctagaagcta gcttttgacc cagagatttg tcagcctgag 1860

accagtatga gatccaaatg ctgcggatct cataagtgat acgaggatca gacaaggtct 1920

ccacccaccg acgaataaag cgttcttgcc tgtctggtgt gaatgagcgg tacctttctc 1980

ctggttgctt gaaattgttc tctttctgaa tgacacactt ctcgcgtttg ccagtgcaca 2040

ttgtagaagg aataggatac ttctcagcat ggcgaacagg atcatacctt gaagggaagt 2100

agtcgatctc ctcatccctg tgcataaaat tcatggagcc atcgtagtga ttgttgtgat 2160

gagcgcattt tggagcatta gcaggtagtt gcaaatagtt tggtccaagt cgatacctct 2220

gggtatcaga gtaggagaaa atacgagttt gaagcatctt atcatctgag taataaaccc 2280

ctggaacaac aatagaaggg cagaaagcta gctgctcatt ctcattagag aagttatcaa 2340

tgttcttgtt cagaactaat cttcccaccg gctgcaaagg caagatatcc tctggccaag 2400

ttttttgtcac atcaagtgga tcaaaatcaa atctgtcttc atgatctgga tccatagtcc 2460

ccgggcagtg ggcgatttga tttaaatctc tagaatagta aattgtaatg ttgtttgttg 2520

tttgttttgt tgtggtaatt gttgtaaaaa tacggatcgt cctgcagtcc tctccaaatg 2580

aaatgaactt ccttatatag aggaagggtc ttgcgaagga tagtgggatt gtgcgtcatc 2640

ccttacgtca gtggagatat cacatcaatc cacttgcttt gaagacgtgg ttggaacgtc 2700

ttctttttcc acgatgctcc tcgtgggtgg gggtccatct ttgggaccac tgtcggcaga 2760

ggcatcttga acgatagcct ttcctttatc gcaatgatgg catttgtagg tgccaccttc 2820

cttttctact gtccttttga tgaagtgaca gatagctggg caatggaatc cgaggaggtt 2880

tcccgatatt acccttttgtt gaaaagtctc aatagccctt tggtcttctg agactgtatc 2940

tttgatattc ttggagtaga cgagagtgtc gtgctccacc atgttgacga agattttctt 3000

cttgtcattg agtcgtaaaa gactctgtat gaactgttcg ccagtcttca cggcgagttc 3060

tgttagatcc tcgatctgaa tttttgactc catggccttt gattcagtag gaactacttt 3120

cttagagact ccaatctcta ttacttgcct tggtttatga agcaagcctt gaatcgtcca 3180

tactggaata gtacttctga tcttgagaaa tatatctttc tctgtgttct tgatgcagtt 3240

agtcctgaat cttttgactg catctttaac cttcttggga aggtatttga tctcctggag 3300

attattactc gggtagatcg tcttgatgag acctgccgcg taggcctctc taaccatctg 3360

tgggtcagca ttctttctga aattgaagag gctaatcttc tcattatcgg tggtgaacat 3420

ggtatcgtca ccttctccgt cgaactttct tcctagatcg tagagataga gaaagtcgtc 3480

catggtgatc tccggggcaa aggagatctc tagagtcgag atttaaatcc taaatcctgc 3540

aggaagctta ccggtataac ttcgtatagc atacattata cgaagttatc catggagcca 3600

tttacaattg aatatatcct gccgccgctg ccgctttgca cccggtggag cttgcatgtt 3660

ggtttctacg cagaactgag ccggttaggc agataatttc cattgagaac tgagccatgt 3720

gcaccttccc cccaacacgg tgagcgacgg ggcaacggag tgatccacat gggacttta 3780

aacatcatcc gtcggatggc gttgcgagag aagcagtcga tccgtgagat cagccgacgc 3840

accgggcagg cgcgcaacac gatcgcaaag tatttgaacg caggtacaat cgagccgacg 3900

ttcacggtac cggaacgacc aagcaagcta gcttagtaaa gccctcgcta gattttaatg 3960

cggatgttgc gattacttcg ccaactattg cgataacaag aaaaagccag cctttcatga 4020

tatatctccc aatttgtgta gggcttatta tgcacgctta aaaataataa aagcagactt 4080

gacctgatag tttggctgtg agcaattatg tgcttagtgc atctaacgct tgagttaagc 4140

cgcgccgcga agcggcgtcg gcttgaacga attgttagac attatttgcc gactaccttg 4200

gtgatctcgc ctttcacgta gtggacaaat tcttccaact gatctgcgcg cgaggccaag 4260

cgatcttctt cttgtccaag ataagcctgt ctagcttcaa gtatgacggg ctgatactgg 4320

gccggcaggc gctccattgc ccagtcggca gcgacatcct tcggcgcgat tttgccggtt 4380

actgcgctgt accaaatgcg ggacaacgta agcactacat ttcgctcatc gccagcccag 4440

tcgggcggcg agttccatag cgttaaggtt tcatttagcg cctcaaatag atcctgttca 4500

ggaaccggat caaagagttc ctccgccgct ggacctacca aggcaacgct atgttctctt 4560

gcttttgtca gcaagatagc cagatcaatg tcgatcgtgg ctggctcgaa gatacctgca 4620

agaatgtcat tgcgctgcca ttctccaaat tgcagttcgc gcttagctgg ataacgccac 4680

ggaatgatgt cgtcgtgcac aacaatggtg acttctacag cgcggagaat ctcgctctct 4740

ccagggggaag ccgaagtttc caaaaggtcg ttgatcaaag ctcgccgcgt tgtttcatca 4800

agccttacgg tcaccgtaac cagcaaatca atatcactgt gtggcttcag gccgccatcc 4860

actgcggagc cgtacaaatg tacggccagc aacgtcggtt cgagatggcg ctcgatgacg 4920

ccaactacct ctgatagttg agtcgatact tcggcgatca ccgcttccct catgatgttt 4980

aactttgttt tagggcgact gccctgctgc gtaacatcgt tgctgctcca taacatcaaa 5040

catcgaccca cggcgtaacg cgcttgctgc ttggatgccc gaggcataga ctgtacccca 5100

aaaaaacagt cataacaagc catgaaaacc gccactgcgc cgttaccacc gctgcgttcg 5160

gtcaaggttc tggaccagtt gcgtgagcgc atacgctact tgcattacag cttacgaacc 5220

gaacaggctt atgtccactg ggttcgtgcc ttcatccgtt tccacggtgt gcgtcacccg 5280

gcaaccttgg gcagcagcga agtcgaggca tttctgtcct ggctggcgaa cgagcgcaag 5340

gtttcggtct ccacgcatcg tcaggcattg gcggccttgc tgttcttcta cggcaagtgc 5400

tgtgcacgga tctgccctgg cttcaggaga tcggaagacc tcggccgtcc gggcgcttgc 5460

cggtggtgct gaccccggat gaagtggttc gcatcctcgg ttttctggaa ggcgagcatc 5520

gtttgttcgc ccagcttctg tatggaacgg gcatgcggat cagtgagggt ttgcaactgc 5580

gggtcaagga tctggatttc gatcacggca cgatcatcgt gcgggagggc aagggctcca 5640

aggatcgggc cttgatgtta cccgagagct tggcacccag cctgcgcgag cagggatcga 5700

tccaacccct ccgctgctat agtgcagtcg gcttctgacg ttcagtgcag ccgtcttctg 5760

aaaacgacat gtcgcacaag tcctaagtta cgcgacaggc tgccgccctg ccctttttcct 5820

ggcgttttct tgtcgcgtgt tttagtcgca taaagtagaa tacttgcgac tagaaccgga 5880

gacattacgc catgaacaag agcgccgccg ctggcctgct gggctatgcc cgcgtcagca 5940

ccgacgacca ggacttgacc aaccaacggg ccgaactgca cgcggccggc tgcaccaagc 6000

tgttttccga gaagatcacc ggcaccaggc gcgaccgccc ggagctggcc aggatgcttg 6060

accacctacg ccctggcgac gttgtgacag tgaccaggct agaccgcctg gcccgcagca 6120

cccgcgacct actggacatt gccgagcgca tccaggaggc cggcgcgggc ctgcgtagcc 6180

tggcagagcc gtgggccgac accaccacgc cggccggccg catggtgttg accgtgttcg 6240

ccggcattgc cgagttcgag cgttccctaa tcatcgaccg cacccggagc gggcgcgagg 6300

ccgccaaggc ccgaggcgtg aagtttggcc cccgccctac cctcaccccg gcacagatcg 6360

cgcacgcccg cgagctgatc gaccaggaag gccgcaccgt gaaagaggcg gctgcactgc 6420

ttggcgtgca tcgctcgacc ctgtaccgcg cacttgagcg cagcgaggaa gtgacgccca 6480

ccgaggccag gcggcgcggt gccttccgtg aggacgcatt gaccgaggcc gacgccctgg 6540

cggccgccga gaatgaacgc caagaggaac aagcatgaaa ccgcaccagg acggccagga 6600

cgaaccgttt ttcattaccg aagagatcga ggcggagatg atcgcggccg ggtacgtgtt 6660

cgagccgccc gcgcacgtct caaccgtgcg gctgcatgaa atcctggccg gtttgtctga 6720

tgccaagctg gcggcctggc cggccagctt ggccgctgaa gaaaccgagc gccgccgtct 6780

aaaaaggtga tgtgtatttg agtaaaacag cttgcgtcat gcggtcgctg cgtatatgat 6840

gcgatgagta aataaacaaa tacgcaaggg gaacgcatga aggttatcgc tgtacttaac 6900

cagaaaggcg ggtcaggcaa gacgaccatc gcaacccatc tagcccgcgc cctgcaactc 6960

gccggggccg atgttctgtt agtcgattcc gatccccagg gcagtgcccg cgattgggcg 7020

gccgtgcggg aagatcaacc gctaaccgtt gtcggcatcg accgcccgac gattgaccgc 7080

gacgtgaagg ccatcggccg gcgcgacttc gtagtgatcg acggagcgcc ccaggcggcg 7140

gacttggctg tgtccgcgat caaggcagcc gacttcgtgc tgattccggt gcagccaagc 7200

ccttacgaca tatgggccac cgccgacctg gtggagctgg ttaagcagcg cattgaggtc 7260

acggatggaa ggctacaagc ggcctttgtc gtgtcgcggg cgatcaaagg cacgcgcatc 7320

ggcggtgagg ttgccgaggc gctggccggg tacgagctgc ccattcttga gtcccgtatc 7380

acgcagcgcg tgagctaccc aggcactgcc gccgccggca caaccgttct tgaatcagaa 7440

cccgagggcg acgctgcccg cgaggtccag gcgctggccg ctgaaattaa atcaaaactc 7500

atttgagtta atgaggtaaa gagaaaatga gcaaaagcac aaacacgcta agtgccggcc 7560

gtccgagcgc acgcagcagc aaggctgcaa cgttggccag cctggcagac acgccagcca 7620

tgaagcgggt caactttcag ttgccggcgg aggatcacac caagctgaag atgtacgcgg 7680

tacgccaagg caagaccatt accgagctgc tatctgaata catcgcgcag ctaccagagt 7740

aaatgagcaa atgaataaat gagtagatga attttagcgg ctaaaggagg cggcatggaa 7800

aatcaagaac aaccaggcac cgacgccgtg gaatgcccca tgtgtggagg aacgggcggt 7860

tggccaggcg taagcggctg ggttgtctgc cggccctgca atggcactgg aacccccaag 7920

cccgaggaat cggcgtgacg gtcgcaaacc atccggcccg gtacaaatcg gcgcggcgct 7980

gggtgatgac ctggtggaga agttgaaggc cgcgcaggcc gcccagcggc aacgcatcga 8040

ggcagaagca cgccccggtg aatcgtggca agcggccgct gatcgaatcc gcaaagaatc 8100

ccggcaaccg ccggcagccg gtgcgccgtc gattaggaag ccgcccaagg gcgacgagca 8160

accagatttt ttcgttccga tgctctatga cgtgggcacc cgcgatagtc gcagcatcat 8220

ggacgtggcc gttttccgtc tgtcgaagcg tgaccgacga gctggcgagg tgatccgcta 8280

cgagcttcca gacgggcacg tagaggtttc cgcagggccg gccggcatgg ccagtgtgtg 8340

ggattacgac ctggtactga tggcggtttc ccatctaacc gaatccatga accgataccg 8400

ggaagggaag ggagacaagc ccggccgcgt gttccgtcca cacgttgcgg acgtactcaa 8460

gttctgccgg cgagccgatg gcggaaagca gaaagacgac ctggtagaaa cctgcattcg 8520

gttaaacacc acgcacgttg ccatgcagcg tacgaagaag gccaagaacg gccgcctggt 8580

gacggtatcc gagggtgaag ccttgattag ccgctacaag atcgtaaaga gcgaaaccgg 8640

gcggccggag tacatcgaga tcgagctagc tgattggatg taccgcgaga tcacagaagg 8700

caagaacccg gacgtgctga cggttcaccc cgattacttt ttgatcgatc ccggcatcgg 8760

ccgttttctc taccgcctgg cacgccgcgc cgcaggcaag gcagaagcca gatggttgtt 8820

caagacgatc tacgaacgca gtggcagcgc cggagagttc aagaagttct gtttcaccgt 8880

gcgcaagctg atcgggtcaa atgacctgcc ggagtacgat ttgaaggagg aggcggggca 8940

ggctggcccg atcctagtca tgcgctaccg caacctgatc gagggcgaag catccgccgg 9000

ttcctaatgt acggagcaga tgctagggca aattgcccta gcaggggaaa aaggtcgaaa 9060

aggtctcttt cctgtggata gcacgtacat tgggaaccca aagccgtaca ttgggaaccg 9120

gaacccgtac attgggaacc caaagccgta cattgggaac cggtcacaca tgtaagtgac 9180

tgatataaaa gagaaaaaag gcgatttttc cgcctaaaac tctttaaaac ttattaaaac 9240

tcttaaaacc cgcctggcct gtgcataact gtctggccag cgcacagccg aagagctgca 9300

aaaagcgcct acccttcggt cgctgcgctc cctacgcccc gccgcttcgc gtcggcctat 9360

cgcggccgct ggccgctcaa aaatggctgg cctacggcca ggcaatctac cagggcgcgg 9420

acaagccgcg ccgtcgccac tcgaccgccg gcgcccacat caaggcaccc tgcctcgcgc 9480

gtttcggtga tgacggtgaa aacctctgac acatgcagct cccggagacg gtcacagctt 9540

gtctgtaagc ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg ggtgttggcg 9600

ggtgtcgggg cgcagccatg acccagtcac gtagcgatag cggagtgtat actggcttaa 9660

ctatgcggca tcagagcaga ttgtactgag agtgcaccat atgcggtgtg aaataccgca 9720

cagatgcgta aggagaaaat accgcatcag gcgctcttcc gcttcctcgc tcactgactc 9780

gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg 9840

gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa 9900

ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc gccccccctga 9960

cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag 10020

ataccaggcg tttccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct 10080

taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc atagctcacg 10140

ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc 10200

ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt 10260

aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta 10320

tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac 10380

agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc 10440

ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat 10500

tacgcgcaga aaaaaaggat ctcaagaaga tccggaaaac gcaagcgcaa agagaaagca 10560

ggtagcttgc agtgggctta catggcgata gctagactgg gcggttttat ggacagcaag 10620

cgaaccggaa ttgcc 10635